Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/AU05/000276

International filing date: 01 March 2005 (01.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: AU
Number: 2004901046
Filing date: 01 March 2004 (01.03.2004)

Date of receipt at the International Bureau: 15 March 2005 (15.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

PCT/AU2005/000276

Australian Government

Patent Office
Canberra

I, LEANNE MYNOTT, MANAGER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2004901046 for a patent by MICHAEL MAN HO MAK as filed on 01 March 2004.

WITNESS my hand this
Tenth day of March 2005

LEANNE MYNOTT
MANAGER EXAMINATION SUPPORT
AND SALES

TITLE

A system for encoding and decoding mobile-phone based ticket-codes using alphanumeric data.

ABSTRACT

A system for encoding and decoding mobile-phone based ticket-codes using alphanumeric data. The invention uses a unique combination of alphanumeric characters that are displayed on the screen of a mobile device to communicate unique information to the handset. This is a superior substitute for traditional one-dimensional and two-dimensional mobile barcodes in many instances, as those codes are not supported by 100% of mobile handsets in the market, and will not be in the foreseeable future.

Many organizations worldwide are beginning to realize the potential benefits of mobile ticketing and mobile voucher dispatch, as it can greatly reduce the cost of paper and post-based ticketing. Mobile tickets and vouchers that do not have a barcode component are easy to fake, but at the same time the graphical capability that barcoding requires is not supported by 100% of the consumer mobile devices in the market.

Alphanumeric coding is a unique string of characters that are essentially encoded and encrypted from the original serial code of the ticket or voucher into a unique alphanumeric string. These original serial codes are often generated from existing ticketing systems. Alphanumeric string has the best of both worlds by supporting 100% of SMS-enabled handsets, but avoids faking with a non-copyable and non-guessable unique code.

This invention deals with the unique encoding and decoding algorithm that will make the system of using alphanumeric code scannable and implementable in a commercial deployment, as well as the unique system of using mobile alphanumeric coding to achieve the business objectives of mobile barcoding in a 100% reliable and robust way. It does it by applying Optical Character Recognition (OCR) in a unique way to be able to optically scan and then algorithmically decode the alphanumeric string reliably back to original serial code.

SUMMARY OF INVENTION

This particular invention deals with:

1. The method of sending unique code data from a business system to a mobile device using alphanumeric string text data;

2. The method of making alphanumeric string data displayed on mobile phones and mobile devices scannable optically;

3. The method of encoding the alphanumeric string so that it is not copyable and guessable, and resilient in the event of scanning errors and missing graphical scan data;

4. The method of decoding a graphical scanned image of an alphanumeric string, to ensure near 100% accuracy;

The invention comprises of the following systems:

5. A method and system that utilise software with the algorithm described by this invention to encode a serial ticket code into an alphanumeric string, which is subsequently sent to a mobile phone using industry-standard communication protocols such as SMPP, that is subsequently scanned by an image capture device such as a digital camera (Refer Diagram 1), and decoded using optical character recognition algorithms, and then decoded again using the algorithm described by this invention, so that at the physical location of the transaction, the original serial code can be obtained with a high-degree of accuracy (99.99% of the time) for validation for a variety of business requirements (Refer Diagram 2);

6. A system that addresses Point 1 by encoding a serial ticket code such as a 16-digit decimal code, into binary bit data. This binary bit data is then applied redundancy algorithm that uses methods such as mirror bits, redundant bits and parity bits to provide resilience of data accuracy. This processed bit data is then encoded into an alphanumeric string using a bit-character mapping system such as one described in Point 7. This alphanumeric code is then encoded again to maximize optical scannability using methods such as one described in Point 8;

7. A system that maps bit data into characters so that bit data with a high number of digits can be encoded into a string short enough for transmission to a mobile device. For example, a 100-bit string can be encoded into a 20-character string using a 5-bit-to-1-character mapping regime. Whilst this invention does not limit the mapping regime to a 5-bit-to-1-character regime, from experimental data, a 5-bit-to-1 regime is recommended for this invention. There are 32 useful characters in this scanning regime, namely the 26 upper-case alphabets, 10 decimal digits, minus zero – that looks like an O, one – that looks like an I or an L, two – that looks like a Z, and eight, that looks like a B. Any other ambiguous alphabetical character can be replaced by its lowercase to maximize accuracy. These 32 characters are then mapped to 5-bit binary data;

8. A system that maximizes scanning accuracy and speed that will utilize a number of methods described below:

    a. Adding a row of equal (=) sign, or any easy to recognize geometry, so that the OCR software can find the orientation of the graphical image quickly and accurately, to enable rapid scanning of the alphanumeric text data;

    b. Adding checkpoint characters in the string using easy to recognize geometries (such as X or O or space or line feed). In the event that the OCR software fails to recognize a character, and miss it in the reported text string, rather than reporting a failed scan, these checkpoints will enable the system to know where to pad the string for missing characters so that the resultant alphanumeric string will have the same length as the original, and with a maximize number of characters in their intend positions;

    c. Giving explicit directional instructions on the scanning unit so that the person operating the scanning unit knows exactly how to place the phone on the scanner for scanning. (eg. “Put your phone here, top side up, within this square”)

9. A system addresses point 2 by using individual and a combination of redundancy bit-based algorithms to create ample data resiliency in the encoded word, optical scan facilitation systems such as one described in Point 8, as well as an iterative scanning system described in point 12;

10. A system that addresses Point 3 by utilising the following methods and systems:

    a. Preservation of the uniqueness of the original serial code;

    b. Usage of random generated codes, and keeping a list of these random generated codes that are valid and/or issued, so that no one can forward guess what the next batch of valid codes will be;

    c. By allowing each code to be only used once, so that it cannot be forwarded or copied for other people's use;

    d. The use of internal checksums and parity bits to detect illegitimate codes or combination of characters;

    e. Redundancy, resilient and geometric regimes described in Point 6, Point 7, Point 8 and Point 9 to provide maximum accuracy;

11. A system that addresses Point 4, by using an image capture device, that will be triggered on by an image detect device, such as an infrared detector. Once turned on, the image capture device (eg. Digital camera) will take a photo of the scanned image. This can be enhanced with the activation of a flash light. Once a photo image is obtained, it will be decoded using optical character recognition algorithms. This will convert the image into text data. This text data is will have missing characters padded with padding characters, and have checkpoint characters removed to give the best guess at the original character string. This string will then be converted into a binary bit string, and have the redundancy algorithm applied to arrive at the original binary bit string.

    Using internal checksums, the redundancy algorithm will ensure that unless it is 99.99% (accuracy threshold – can be set different) sure that the resultant binary bit string is same as the original one, the check will fail, and the scan will be forced to be done again by system described in Point 12.

    This original binary string is converted back to the original serial code, for validation, reporting, transaction authorization or for whatever purpose the business requirement prescribes. Refer to Diagram 2 for this process.

12. Due to fluctuations in lighting, positional, device-specific and optical factors, a scan may fail by not having sufficient accuracy for the redundancy algorithms to return the original binary string with the confidence level sufficient for the accuracy threshold. In this instance, the system in Point 11 will be forced to repeat the scanning process. This process will keep repeating itself until the scan succeeds. The person performing the scan will keep attempting to scan the same mobile device, because he/she will not receive any positive feedback from the system (such as an audio or video signal) until this succeeds. At some stage the user may give up on that particular code or phone, which suggests that the code is not scannable, or it is not valid.

DIAGRAM 1

ENCODING & DISPATCHING

1. Original paper-based ticket in 16-digit decimal serial code format

1456722135924227

2. Convert to 60-bit binary bit format to prepare for data encoding

0010101001010111
1010101010100101
0101101101010010..

3. Apply redundancy algorithms to provide resilience and error correction, and resultant code can be a 100-bit code

1110010110101011
1101010101010010
1010110110101001
0010101001010111
1010101010100....

4. Encode resultant bit code into 20-digit alphanumeric string using 32-bit alphanumeric code mapping

FK45K324JK
HY7294FFPZ

5. Add data checkpoints and reference geometries to maximize speed and accuracy of scanning

OFK4O5K3
O24JKOHY7
O294OFFPZ

6. The alphanumeric string is optically scanned and converted into text

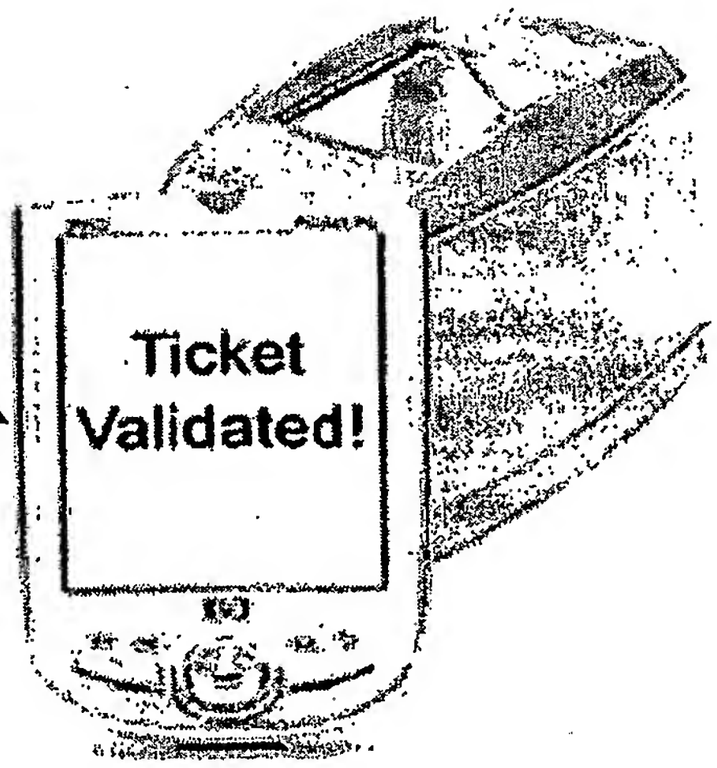
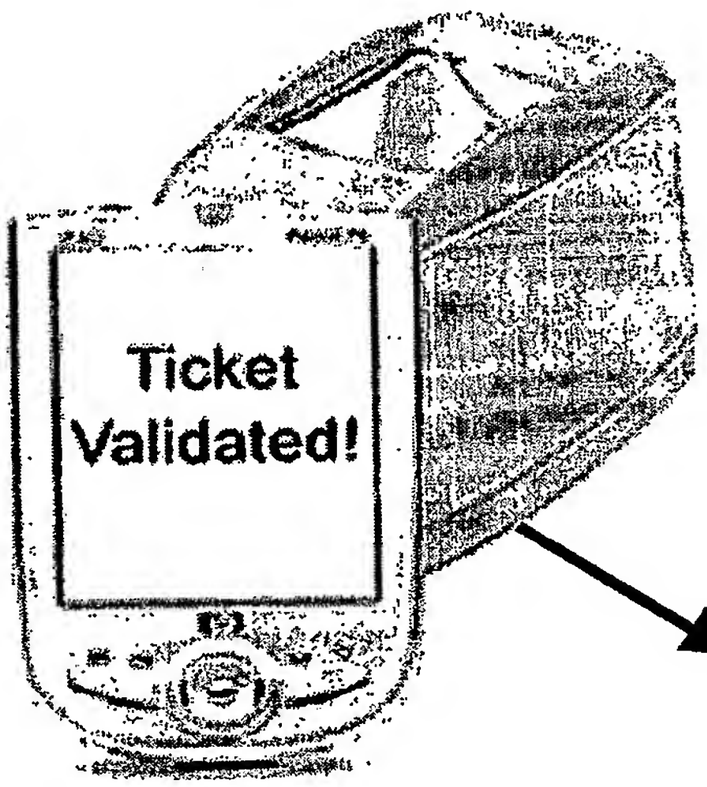


DIAGRAM 2

DECODING



1. Once the alphanumeric string is scanned, the OCR software will attempt to decode the graphical scanned image back into text

2. The output of the OCR algorithm will be a string of text. This string of text may contain defects such as missing or wrong characters

OFK4O6K3
O24JKOHY7
O2_4OFEPZ

3. After padding for missing characters using checkpoint information, it is converted back to a 100-digit bit string

1010010110101011
1101010101011110
1010110111010010
0100000101011110
101010101000....

4. Inverse redundancy algorithm will arrive at the original 60-digit without any errors or inconsistency. Or the scan will fail and attempt again

0010101001010111
1010101010100101
0101101101010 01..

1456722135924227

5. Converted back into original code for validation with internal or external infrastructure